APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/965,052	09/25/2001	Michael J. Payne	42390P11780	2670

		EXAMINER
7590	10/04/2004	SKED, MATTHEW J

James H. Salter
Blakely, Sokoloff, Taylor & Zafman
Seventh Floor
12400 Wilshire Boulevard
Los Angeles, CA 90025-1030

ART UNIT	PAPER NUMBER
2655	

DATE MAILED: 10/04/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

-- *The MAILING DATE of this communication appears on the cover sheet with the correspondence address --*

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE <u>3</u> MONTH(S) FROM
THE MAILING DATE OF THIS COMMUNICATION.
- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed
  after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
  Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any
  earned patent term adjustment. See 37 CFR 1.704(b).

Status

1) ☐ Responsive to communication(s) filed on _____ .

2a) ☐ This action is **FINAL.**   2b) ☒ This action is non-final.

3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is
closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4) ☒ Claim(s) *1-47* is/are pending in the application.

    4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5) ☐ Claim(s) _____ is/are allowed.

6) ☒ Claim(s) *1-47* is/are rejected.

7) ☐ Claim(s) _____ is/are objected to.

8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9) ☒ The specification is objected to by the Examiner.

10) ☒ The drawing(s) filed on <u>25 September 2001</u> is/are: a) ☐ accepted or b) ☒ objected to by the Examiner.

    Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

    Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

    a) ☐ All   b) ☐ Some * c) ☐ None of:

      1. ☐ Certified copies of the priority documents have been received.

      2. ☐ Certified copies of the priority documents have been received in Application No. _____ .

      3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage
application from the International Bureau (PCT Rule 17.2(a)).

    * See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)

2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)
Paper No(s)/Mail Date *5/8/03 and 3/8/04.*

4) ☐ Interview Summary (PTO-413)
    Paper No(s)/Mail Date. _____ .

5) ☐ Notice of Informal Patent Application (PTO-152)

6) ☐ Other: _____ .

DETAILED ACTION

Drawings

1.   The drawings are objected to because Figures 1 and 2 are sloppy. Corrected drawing sheets in compliance with 37 CFR 1.121(d) are required in reply to the Office action to avoid abandonment of the application. The objection to the drawings will not be held in abeyance.

Specification

2.   The disclosure and claims 8 and 23 are objected to because the term "voice recognition" is misused for what nowadays is called --**speech recognition**-- in the speech signal processing art. While "voice recognition" and "speech recognition" were both once used interchangeably to refer to spoken word recognition, nowadays these two terms are distinguished. The term "**voice** recognition" now denotes identification of *who* is doing the speaking (class 704/246), while "**speech** recognition" (or "**word** recognition") denotes identification of *what* is being said (class 704/251). So, appropriate correction to the proper terms of art is required.

Claim Rejections - 35 USC § 102

3.   The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

4.   Claims 13, 14, 30, 31, 40, and 41 are rejected under 35 U.S.C. 102(b) as being

anticipated by Stanford et al. (U.S. Pat. 5,615,296).

Regarding claims 13, 30, and 40, Stanford teaches a speech translation method,

apparatus and computer readable medium containing executable computer program

instructions comprising:

generating a first phoneme from a first audio signal using a first context of a

language vocabulary (uses first context to give user available speech options and

switches to second context once the first audio signal is determined, col. 10, lines 65-67

and col. 11, lines 1-8);

switching said first context to a second context (switches to types of restaurants

after selection, col. 11, lines 3-8); and

generating a second phoneme from a second audio signal using said second

context of the language vocabulary (col. 11, lines 8-11).

5.   As per claims 14 and 31, Stanford suggests the real-time speech translation is

maintained (context switching is near-instantaneous hence the system will continue to

operate in real-time as a singular vocabulary speech translation device, col. 14, lines

10-16).

6.   As per claim 41, Stanford teaches the apparatus being a business entity being a

restaurant services company (restaurant help desk, col. 11, lines 11-16).

Claim Rejections - 35 USC § 103

7.   The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all

obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

8.   Claims 1, 5, 18, 22, 35-39, 44 and 46-47 are rejected under 35 U.S.C. 103(a) as

being unpatentable over Stanford in view of Alshawi et al. (U.S. Pat.

Pub2002/0072914A1).

Regarding claims 1, 18, 35, 44 and 46, Stanford teaches a speech translation

method, apparatus and computer readable medium containing executable computer

program instructions comprising:

limiting the language vocabulary to a subset of the language vocabulary (user

chooses the type of movie to access that subset of the vocabulary, col. 6, lines 40-43);

separating said subset into at least two contexts (recent releases and all-time

hits, col. 6, lines 48-50); and

associating the speech signal with at least one of said at least two contexts and

performing speech recognition using the context (user selects recent releases and new

choices are displayed, col. 6, lines 52-54).

Stanford does not teach the translating the speech signal into text.

Alshawi teaches a method for translation speech into text that uses multiple

contexts within the vocabulary (converts audio to text form, paragraph 30).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Stanford to translate the speech signal into text as taught by Alshawi because it would allow the user to visualize the recognition and hence allow the user to verify that the recognition was successful.

9.   Regarding claims 5 and 22, Stanford and Alshawi do not teach that the subset is selected from a group consisting of a medical subset.

However, the Examiner takes Official Notice that medical vocabularies are common in speech recognition and it would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Stanford and Alshawi to teach that the subset is selected from a group consisting of a medical subset because having a medical vocabulary would facilitate a hands free speech commanded device which would be beneficial to those in the medical field.

10.   As per claims 45 and 47, Stanford teaches the apparatus being a business entity being a restaurant services company (restaurant help desk, col. 11, lines 11-16).

11.   Claims 2-4 and 19-21 are rejected under 35 U.S.C. 103(a) as being unpatentable over Stanford in view of Alshawi taken in further view of Knittle (U.S. Pat. 5,758,319).

As per claims 2 and 19, Neither Stanford or Alshawi teach applying a constraint filter to at least one context of said a least two contexts to restrict the size of said subset associated with said at least one context.

Knittle teaches a method for limiting the number of words searched by a speech recognition program by using a sub-vocabulary generator that identifies specific words within the vocabulary (col. 3, lines 27-33).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Stanford and Alshawi to apply a constraint filter as taught by Knittle to at least one of the contexts because it would further limit the amount of searching the speech recognition program would perform hence making the system faster.

12.   As per claim 3 and 20, Stanford, Alshawi and Knittle do not teach that the constraint filter is at least one of a set of patients and a set of frequently prescribed drugs.

However, the Examiner takes Official Notice that speech recognition with a medicine-related vocabulary is well known. Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Stanford, Alshawi and Knittle to teach a medicine-related vocabulary having a constraint filter of at least one of a set of patients and a set of frequently prescribed drugs because it would increase recognition speed and accuracy when used in a physician's office.

13.   Regarding claims 4 and 21, neither Stanford nor Alshawi teach that speech recognition is biased using said constraint filter.

Knittle teaches that the recognizer only searches the predetermined portion of the language model hence speech recognition is biased (col. 3, lines 33-36).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Stanford and Alshawi so that speech recognition is biased using said constraint filter as taught by Knittle because it would search only the smaller vocabulary hence speeding up recognition and making it more accurate.

14.   As per claim 36, Stanford teaches a display (kiosk or TV screen, col. 6, lines 42-44).

Stanford does not teach the display to display the resultant text.

Alshawi teaches a speech-to-text system with a display (paragraphs 29 and 30).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Stanford to display the resultant text as taught by Alshawi because it would allow the user to visualize the recognition and hence allow the user to verify that the recognition was successful.

15.   As per claim 37, Stanford does not teach a wireless interface to allow communication of the speech signal.

Alshawi teaches that the system can interface over a cellular or satellite network (paragraph 25).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Stanford to have a wireless interface as taught by Alshawi because it makes the system more flexible hence facilitating the interaction between the user and the device.

16.   As per claim 38, Stanford teaches the apparatus is installed in a vehicle (car, col. 4, lines 48-50).

17.   As per claim 39, Stanford does not teach the apparatus to communicate with the Internet.

Alshawi teaches the apparatus to communicate with the Internet (paragraph 69).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Stanford to communicate with the internet as taught by Alshawi because it would allow the models to be stored on another device hence saving more memory on the apparatus.

18.   Claims 6-8 and 23-25 are rejected under 35 U.S.C. 103(a) as being unpatentable over Stanford in view of Knittle.

As per claims 6 and 23, Stanford teaches a method and computer readable medium containing executable computer program instructions of designing a speaker independent speech recognition speech-enabled user interface comprising:

defining a subject matter to base the user interface on (program for video rentals, col. 6, lines 40-41); and

designating a first allowable vocabulary for a first speech enabled field of the user interface and designation a second allowable vocabulary for a second speech enabled filed of the user interface (multiple fields each with their own vocabulary, col. 6, lines 41-51).

Stanford does not teach designing a constraint filter for at least one of said first allowable vocabulary and second allowable vocabulary.

Knittle teaches a method for limiting the number of words searched by a speech recognition program by using a sub-vocabulary generator that identifies specific words within the vocabulary (col. 3, lines 27-33).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Stanford to apply a constraint filter as taught by Knittle

to at least one of the contexts because it would further limit the amount of searching the speech recognition program would perform hence making the system faster and more accurate.

19.   Regarding claims 7, 8, 24 and 25, neither Stanford nor Knittle teach that the subject matter is a medical subject matter, characterized by at least one of; a medical application, and a medical setting.

However, the Examiner takes Official Notice that medical vocabularies are well known in speech recognition. It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Stanford and Alshawi to teach that the subset is selected from a group consisting of such medical subset because having a medical vocabulary would facilitate a hands free speech commanded device in a physician's office which would be beneficial to those in the medical field.

20.   Claims 9-12 and 26-29 are rejected under 35 U.S.C. 103(a) as being unpatentable over Padmanabhan et al. (U.S. Pat. 6,385,579) in view of Haddock (U.S. Pat. 5,983,187) and taken in further view of Stanford.

As per claims 9 and 26, Padmanabhan teaches a method and computer readable medium containing executable computer program instructions for translating speech signal into text comprising:

identifying and extracting a segment of the audio signal (partitions the signal into frames, col. 5, lines 20-23);

generating sets of phonemes that correspond to the segment of the audio signal (hypothesizes a sequence of words, col. 5, lines 38-40);

rating the sets of phonemes for accuracy as an individual word and as a part of a larger word (average phone recognition probability of the compound word and individual words, col. 9, lines 2-6);

combining accuracy ratings from said rating (acoustic measure represents the difference between the two average phone recognition probabilities, col. 9, lines 2-6); and

selecting the word or part of the word corresponding to the segment of the audio signal (hypothesis with best score is outputted as the recognized sequence, col. 5, lines 40-42).

Padmanabhan does not teach ranking the sets of phonemes according to said rating.

However, the Examiner takes Official Notice that ranking a set of possible word scores is common in the art and it would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Padmanabhan to rank the sets of phonemes according to the rating because it would give the user a ordered look at potential words based upon their score hence facilitating a choice.

Padmanabhan does not teach identifying at least two anchor points in the audio signal wherein the segment of the audio signal is contained between the at least two anchor points.

Haddock teaches a system for keyword recognition that identifies anchor points through silence detection to identify the sections of speech (col. 3, lines 54-59).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Padmanabhan to identify and use anchor points in the audio signal to obtain the speech segments for recognition as taught by Haddock because this enables recognition of keywords or phrases of interest.

Neither Padmanabhan nor Haddock teach generating sets of phonemes using a subset of a language vocabulary.

Stanford teaches partitioning the language vocabulary into subsets (user chooses the type of movie to access that subset of the vocabulary, col. 6, lines 40-43).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Padmanabhan and Haddock to generate phonemes using a subset of a language vocabulary as taught by Stanford because it would give a more concentrated vocabulary and therefore speeding up searching time.

21.   As per claims 10 and 27, neither Padmanabhan nor Haddock teach the subset of the language vocabulary is separated into a plurality of contexts and said generating is performed within a context of the plurality of contexts.

Stanford teaches that the subset of the language vocabulary is separated into a plurality of contexts (recent releases and all-time hits, col. 6, lines 48-50).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Padmanabhan and Haddock to separate the vocabulary subset into a plurality of contexts as taught by Stanford because further limit the amount of terms in the vocabulary hence speeding up searching.

22.   As per claim 11 and 28, neither Padmanabhan nor Haddock teach the context is dynamically changed during generating.

Stanford teaches dynamically changing the context (col. 13, lines 21-24).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Padmanabhan and Haddock to dynamically change the context as taught by Stanford because it would give more flexibility by allowing multiple contexts to be used hence allowing better recognition.

23.   As per claim 12 and 29, Padmanabhan does not teach identifying a new anchor point, such that said generating is performed on a segment of the audio signal defined with the new anchor point.

Haddock suggests multiple anchor points and sections of speech corresponding to these anchor points, hence any anchor points found after the initial anchor points would be new (col. 3, lines 54-59).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Padmanabhan to identify a new anchor point to determine a new segment to be processed as taught by Haddock because it would allow many segments of speech to be identified and processed hence allowing the user to speak more naturally.

24.   Claims 15,16, 32 and 33 are rejected under 35 U.S.C. 103(a) as being unpatentable over Walker et al. (U.S. Pat. 6,434,529).

Regarding claims 15 and 32, Walker suggests a speech translation method and computer readable medium containing executable computer program instructions comprising:

generating a first and second phoneme from an audio signal using a first and second context of a language vocabulary (activates multiple grammars, compares the incoming audio to these grammars, result events are returned and these results may include alternative guesses, col. 12, lines 19-34 and col. 17, lines 19-24); and

selecting a word or part of a word from the first phoneme and the second phoneme that represents a translation of the audio signal (when recognition is completed only one result accepted event is provided hence a selection must be made, col. 12, lines 35-38).

25.   As per claims 16 and 33, Walker teaches that real-time speech translation is maintained (col. 14, lines 62-67).

26.   Claims 17 and 34 are rejected under 35 U.S.C. 103(a) as being unpatentable over Walker in view of Stanford.

Walker does not teach that the first context is switched to said second context before said generating the second phoneme.

Stanford teaches dynamically changing the context (col. 13, lines 21-24).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Walker to dynamically change the context as taught by Stanford before generating the second phoneme because it would save memory by allowing only one context in the active memory at a time.

27.   Claims 42 and 43 are rejected under 35 U.S.C. 103(a) as being unpatentable over Walker et al. (U.S. Pat. 6,434,529) in view of Stanford.

Regarding claim 42, Walker suggests an apparatus comprising:

generating a first and second phoneme from an audio signal using a first and second context of a language vocabulary (activates multiple grammars, compares the incoming audio to these grammars, result events are returned and these results may include alternative guesses, col. 12, lines 19-34 and col. 17, lines 19-24).

Walker does not teach that the first context is switched to said second context before said generating the second phoneme.

Stanford teaches dynamically changing the context (col. 13, lines 21-24).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Walker to dynamically change the context as taught by Stanford before generating the second phoneme because it would save memory by allowing only one context in the active memory at a time.

28.   As per claim 43, Walker teaches the apparatus further comprises a business entity being a restaurant services company (system takes and processes food order from user, col. 6, lines 41-54).

Conclusion

29.   The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. Stanford et al. (U.S. Pat. 5,513,298) teaches a system with vocabulary subsets each with multiple contexts. Lee et al. (U.S. Pat. Pub.
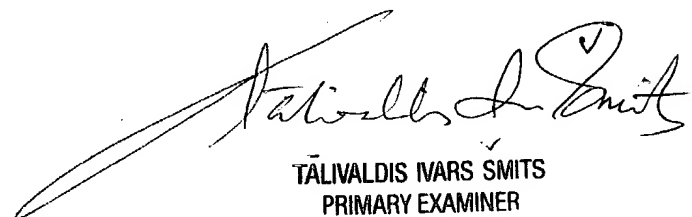
2002/0087313), Raud et al. (U.S. Pat. Pub. 6,125,341), Cohen (U.S. Pat. 6,571,209), and Sabourin et al. (U.S. Pat. 5,987,414) teach a vocabulary with subsets for faster vocabulary searching time during speech recognition.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Matthew J Sked whose telephone number is (703) 305-8663. The examiner can normally be reached on Mon-Fri (8:00 am - 4:30 pm).

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Talivaldis Smits can be reached on (703) 306-3011. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

MS
09/27/04

TALIVALDIS IVARS SMITS
PRIMARY EXAMINER